

Obtaining Data on Patient Race, Ethnicity, and Primary Language in Health Care Organizations: Current Challenges and Proposed Solutions

Romana Hasnain-Wynia and David W. Baker

Objectives. To provide an overview of why health care organizations (HCOs) should collect race, ethnicity, and language data, review current practices, discuss the rationale for collecting this information directly from patients, and describe barriers and solutions.

Principal Findings. Hospitals and HCOs with data from their own institutions may be more likely to look at disparities in care, design targeted programs to improve quality of care, and provide patient-centered care. Yet data collection is fragmented and incomplete within and across organizations. A major factor affecting the quality of data is the lack of understanding about how best to collect this information from patients.

Conclusions. If HCOs make a commitment to systematically collect race/ethnicity and language data from patients, it would be a major step in enhancing the ability of HCOs to monitor health care processes and outcomes for different population groups, target quality initiatives more efficiently and effectively, and provide patient-centered care.

Key Words. disparities, quality of care, patient-centered care, self-reported race/ethnicity, uniform framework, and health care organizations

Numerous studies document that racial and ethnic minorities often receive lower quality care than nonminorities. Although aggregate national data are important, sample sizes often limit their usefulness to only broad racial and ethnic groups. In addition, the data in these surveys may come from records rather than direct interviews of individuals and the information may be based on the observation of the person filling out the record. All these factors leave the quality and consistency of the data questionable. Although much information on health care comes from health care organizations (HCOs) (hospitals, health plans, and medical groups), data on race, ethnicity, and language are often not available or are incomplete (Ver Ploeg and Perrin 2004). In this

paper, we focus on the collection of race, ethnicity, and language data by HCOs.

Valid and reliable data are fundamental building blocks for identifying differences in care and developing targeted interventions to improve the quality of care delivered to specific population groups. There have been clear calls to action to systematically document disparities and tailor interventions to improve the quality of care. In fact, the drive toward measuring quality is based on the idea that performance measures can help patients, consumers, providers, and purchasers understand what high-quality health care is and increase demand for it. The capacity to measure and monitor quality of care for various racial/ethnic populations rests on the ability both to measure quality of care in general and to conduct similar measurements across different racial/ethnic groups (Fremont and Lurie 2004; Lurie, Jung, and Lavizzo-Mourey 2005).

The Institute of Medicine (IOM) *Crossing the Quality Chasm* report focuses on the quality gap, identifies current practices that impede quality care, and explores how systems approaches can be used to implement change. The subsequent IOM report, *Unequal Treatment: Confronting Racial and Ethnic Disparities in Healthcare* places its focus on disparities in health care and concludes that racial and ethnic minorities are less likely to receive even routine medical procedures and experience lower quality health care. Further, these two reports and others urge collecting data on patient race, ethnicity, and language. The report by the National Research Council of the National Academies, *Eliminating Health Disparities: Measurement, and Data Needs*, speaks directly to the importance of collecting valid and reliable data to reduce disparities and improve quality.

Reflecting this mandate, efforts are underway by America's Health Insurance Plans (AHIP) to improve the collection of race, ethnicity, and primary language data in health plans and by the Health Research and Educational Trust (HRET), the research and educational affiliate of the American Hospital Association, to improve data collection in hospitals. The public appears to support collecting this information. Attempts to eliminate the collection of race and ethnicity data in California under Proposition 54 were soundly defeated when opposition arose (Torrassa 2003). A national survey (Robert Wood Johnson Foundation 2003) of adults found that over 50 percent of the

Address correspondence to Romana Hasnain-Wynia, Ph.D., Vice President, Research, Health Research and Educational Trust/AHA, One North Franklin, 30th Floor, Chicago, IL 60606. David W. Baker, M.D., M.P.H., FACP, is Chief of the Division of General Internal Medicine, Feinberg School of Medicine of Northwestern University, Chicago, IL.

respondents favor legislation allowing race/ethnicity data collection when told of its benefits.

A recent study conducted by the authors (Baker et al. 2005, 2006) of patients' attitudes towards health care providers collecting information about their race and ethnicity found that 80 percent agreed that health care providers should collect information on patients' race and ethnicity, but many felt uncomfortable giving this information. We discuss findings from this study in more detail in the section entitled "Barriers to Collecting Data Directly from Patients."

In this article, we provide an overview of why HCOs should collect race, ethnicity, and language data and review current practices. We discuss the rationale for collecting this information directly from patients and/or enrollees (i.e., self-report), describe foreseeable obstacles, and explicate the mechanics, even the art, of overcoming them. We propose recommendations for standardizing data collection practices and discuss policy implications.

WHY COLLECT RACE, ETHNICITY, AND PRIMARY LANGUAGE DATA IN HCOs

HCOs should collect information on patients' race, ethnicity, and language so that they can understand the needs of the population they serve, measure disparities in care within their institution, initiate programs to improve quality of care, and provide patient-centered care. Aaron and Clancy (2003) assert that a growing consensus accepts a strategy integrating reduction in disparities in quality of care as a coherent and efficient approach to redesigning the U.S. health care system. David Williams spoke of the inherent tension in blending the fields of quality and healthcare disparities at the Robert Wood Johnson Foundation Conference on Disparities and Quality of Care (2005) stating, "...we cannot necessarily assume that improving quality will reduce disparities, therefore race-specific strategies may be necessary," whereas Judith Hibbard stated, "...we can improve quality by providing patient-centered care." Both tactics require HCOs to know who their patients are.

Responding to Communities and Providing Patient-Centered Care

Communities want HCOs to be accountable and responsive to them. According to the American College of Physicians (ACP) position paper on racial and ethnic disparities in health care "an ongoing dialogue with surrounding communities can help a HCO integrate cultural beliefs and perspectives into

health care practices and health promotion activities”(2004). Further, effectively managing the health care of patients and using culturally appropriate care improves the health of communities (ACP 2004).

Tracking the racial/ethnic and language composition with concurrently changing health care needs of communities is vital if HCOs are to fulfill their functions. Race/ethnicity and linguistic norms contribute to the definition of what symptoms are noteworthy and how symptoms are presented (Bartlett et al. 1984). As HCOs strive to develop systems and practices that ensure care of all patients is truly patient-centered, understanding the racial/ethnic and language contexts of their patient population is vital.

Accurate information on patients' race, ethnicity, and language is essential to ensure the adequacy of interpreter services, patient information materials, and cultural competency training for staff. Ideally, this information can also be linked to quality measures to examine disparities and undertake targeted quality improvement programs to eliminate disparities. For example, a study of 58,700 randomly selected hemodialysis patients showed that improving the process of care and targeted quality improvement efforts can reduce disparities (Sehgal 2003). However, it was also notable in this study that there was a reduction in disparities in two of the measures that mainly reflected changes in physician behavior. However, this reduction in race disparity was not seen in another measure that was more dependent on patient behavior. This finding underscores the importance of self-identified race/ethnicity and the unique cultural and behavior aspects related to it. This information can help HCOs identify and understand when to implement more generic quality improvement initiatives versus more culturally appropriate and targeted quality improvement interventions to specific patient subgroups.

Mandates and Accreditation

External directives, ranging from federal and state reporting requirements to accreditation, also require the collection of these data. The Center for Medicare and Medicaid Services (CMS) has implemented policies to use race and ethnicity data for quality improvement purposes. For example, peer review organizations in all 50 states contract with CMS to assess and promote quality of health services and have been charged with reducing disparities. Under the Medicare+Choice Quality Assessment Performance Improvement project, managed care plans are required to identify racial and ethnic disparities in clinical outcomes. Some state strategies for Medicaid managed care quality assessment and improvement ask for the inclusion of race, ethnicity, and language data (Perot and Youdelman 2001).

On the accreditation front, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) has field-tested a new standard for collecting race, ethnicity, and language data. At this time, the National Committee for Quality Assurance (NCQA) does not require reporting on race and ethnicity, but it has convened an Expert Advisory Panel to assess the feasibility of developing new standards for the collection of race and ethnicity data by health plans (Donohue 2004). In the future, hospitals, clinics, medical groups, and health plans will likely have a clear line of accountability that include expectations about the absence of disparities and inconsistencies in care.

Quality Improvement Efforts

Many experts have called for HCOs to stratify their quality reports by race and ethnicity (Fiscella et al. 2000). Routine reporting of racial and ethnic disparities would highlight existing problems and studies suggest that public reporting of quality data does, in fact, motivate organizations to develop and implement quality improvement efforts (Marshall et al. 2000). As part of the Hospital Quality Alliance, over 4,000 hospitals are voluntarily reporting inpatient quality of care measures to CMS. With the increasing national focus on racial and ethnic disparities as a quality of care problem, it seems likely that CMS will take the Hospital Quality Alliance to the next logical step and ask hospitals to stratify their quality measures by race and ethnicity. The Department of Health and Human Services Committee on the Collection of Race and Ethnicity data recommends that measures of race and ethnicity be obtained in all health care data systems and that linkages of data be used whenever possible (National Committee on Vital and Health Statistics 2004). More recently, Senators Lieberman (D-Conn) and Hatch (R-Utah) have proposed legislation requiring hospitals to collect patients' race information and have tied it with higher Medicare payments (Modernhealthcare.com 2005). Private foundations recognize the importance of this issue. The Commonwealth Fund has funded two studies to examine linking race and ethnicity data to inpatient quality of care measures. The Robert Wood Johnson Foundation's national initiative to address disparities in cardiac care, *Expecting Success: Excellence in Cardiac Care*, focuses reporting cardiac care quality measures by race/ethnicity and language.

CURRENT PRACTICES

Data collection practices for race, ethnicity, and primary language information are quite variable, sources are diverse and fragmented, and the information is

incomplete (Melnick and Perrin 2003). There is a prevailing misperception within some HCOs that it is illegal to collect race and ethnicity information from patients. Federal statutes do not prohibit collecting these data, and a few even require it. Indeed, collecting and reporting data on race, ethnicity, and language are legal and authorized under Title VI of the Civil Rights Act of 1964 (Perot and Youdelman 2001).

Hospitals

Hospitals play a major role in a community's health care delivery system. As communities become more diverse, hospitals are challenged to design and implement programs and treatment protocols that reduce disparities and improve quality of care (Ver Ploeg and Perrin 2004). However, the infrastructure for collecting and using race, ethnicity, and language data in hospitals is underdeveloped leading to problems of redundancy, inefficiency, and inaccuracy.

A Commonwealth Fund Report (Hasnain-Wynia, Pittman, and Pierce 2004) focusing on the current state of race, ethnicity, and primary language data collection in hospitals indicates that 78 percent of hospitals collect information on patient race and ethnicity, but the quality of these data is quite poor. Of these, 56 percent indicated that more than one unit or clinic within the hospital collected these data, but that this information was not shared. Despite expert recommendations that patients be asked to self-report their race and ethnicity, over half of the hospitals obtained these data by observation of admitting or registration staff, especially in the emergency department. Most (86 percent) hospitals provide a limited number of categories that patients or guardians can select to indicate race and ethnicity; 13 percent give patients the option of using a "fill in the blank" response, though often these text responses are not coded. A small percentage of hospitals (10 percent) provide more granular categories tailored to their local patient/community demographics, suggesting that some hospitals are willing to adjust categories to capture information on important patient subgroups.

Medical Group Practices

Little is known about the collection of data on race and ethnicity in medical group practices. Medical groups are less likely than hospitals to collect race and ethnicity data (Nerenz, Currier, and Paez 2004). Seventy-five percent of medical groups that responded to one survey did not collect race/ethnicity data because they thought it was unnecessary or that collection was potentially disturbing to patients. Medical groups that collected the data did so primarily

for internal quality improvement or disease management purposes, and some were closely affiliated with hospitals that collected data on race/ethnicity as part of the inpatient registration process.

This is a particularly important gap. We know far less about health care disparities in the outpatient as compared with the inpatient setting. To some degree, this reflects a general lack of knowledge about quality of care in the outpatient setting. However, the increasing adoption of electronic medical records systems should help surmount this barrier, and routine data collection by race/ethnicity in medical groups would greatly expand our knowledge and help identify targets for interventions.

Community Health Centers (CHCs)

Perhaps because of the tremendous diversity of patients seen at CHCs and their mandate to have their boards of directors represent their communities, CHCs appear to be ahead of the curve in collecting information on patient race and ethnicity. The Bureau of Primary Health Care (BPHC) has demonstrated success in collecting data in this setting. BPHC's Universal Data System stores data from 700 grantees at 3,000 health care sites. With the backing of a statutory mandate, the BPHC has established specific racial, ethnic, and primary language data collection and reporting requirements applicable to its network of CHCs. Enrollment data have been collected periodically, and plans are underway to secure disaggregated data to assess and report on clinical outcomes by race, ethnicity, and primary language.

There are three specific statutes that explicitly require the collection of demographic information, including race and ethnicity (Perot and Youdelman 2001). These include: (1) Grantees of the Maternal and Child Health (MCH) Services Block Grant are required to provide annual deliveries broken down by racial and ethnic group and the number of women who were provided prenatal, delivery, or postpartum care under MCH or Medicaid; (2) Evaluation of the Substance Abuse and Mental Health Services Administration (SAMHSA) mental health services is broken down by race and ethnicity; and (3) Grantees of SAMSHA who provide services to children of substance abusers are required to collect data on the ethnicity of the children served. The success of these programs provides evidence to other medical groups that this information can be routinely obtained from patients in outpatient practice.

Health Plans

In a study released by America's Health Insurance Plans (AHIP) and the Robert Wood Johnson Foundation, over half of the enrollees in 137 plans that

responded to a survey are enrolled in plans that collect data on race and ethnicity (2004). The majority of plans in the study (74.1 percent) collect this information at enrollment, but 40 percent indicated that they obtain this information through files linked to external data sources using geocoding software or other proxies such as links to federal agencies or birth records (America's Health Insurance Plans 2004). An earlier preliminary study conducted by Nerenz et al. (2002) had found that health plans generally did not routinely capture information on the race and ethnicity of their members and did not analyze quality of care for members of specific racial and ethnic groups. The collection of race and ethnicity data by health plans is inconsistent across the industry, but health plans generally are supportive of collecting these data (Bocchino 2004).

SELF-REPORT IS MORE ACCURATE THAN STAFF OBSERVATIONS

Although many studies have been conducted to examine differences in care for whites and blacks, far less is known about quality of care for Latinos, Asian Americans, or Native Americans. Even less is known about disparities within subpopulations of these large groups (e.g., Mexican versus Puerto Rican, or Korean versus Chinese) and which types of quality improvement interventions work at improving care (AHRQ, U.S. Department of Health and Human Services 2003). If HCOs are committed to providing patient-centered care, reducing disparities, and improving the quality of care, they will need to collect information about their patients' race/ethnicity and language, which will need to come directly from patients.

Having a clerk attempt to classify patients based upon their appearances may be an unfortunate legacy of past times when, as in the first U.S. population census in 1790, we defined people as "whites" or "others." It was even somewhat controversial that the 2000 Census incorporated 15 racial categories, including write-in "other" options, and for the first time allowed respondents to identify themselves as having a multiracial or multiethnic background. It is possible to move beyond the broad historical categories that have been used to categorize individuals and obtain more nuanced information if we collect this information through self-report.

Self-Report Is More Accurate

One of the most important rationales for collecting race and ethnicity information directly from patients is that you get inaccurate information if you do

not ask. Third parties are not the best judges of race, ethnicity, and language (Hasnain-Wynia, Pittman, and Pierce 2004). Research has shown that most observers (e.g., admissions or registration clerks) will accurately identify individuals as white or black, but multiracial and Hispanic individuals are often misidentified (Smith 1997). In a study examining validity of race and ethnicity classifications, respondents described themselves in ways that were inconsistent with categories included in a hospital registration database and many respondents were assigned to categories that were inconsistent with their self-reported identities (Moscou et al. 2003). This is particularly problematic for dark-skinned individuals from Caribbean countries who may identify themselves as Hispanic but be classified by an observer as black.

The lack of reliability in not asking directly is further compounded for individuals with multiracial/ethnic backgrounds. Offering fixed categories to choose from is not a much better solution because it assumes individuals can pigeonhole themselves into one category. Not only can this prospect be confusing, but also it may well be insulting to patients who want to fully recognize their multiethnic/racial heritage and not be classified as “other.”

Self-Identification Reveals More than Broad Categories

If someone identifies herself as El Salvadoran, this tells a great deal more than if she checks a box saying “Hispanic/Latina.” She is saying, “I view myself as part of the El Salvadoran community.” In many cities, there is no “Latino” community or “Asian American” community: there are communities of Mexicans, Puerto Ricans, Cubans, Chinese, and Koreans. Individuals who checked “other” on the 1990 U.S. Census chose specifications such as Puerto Rican or Panamanian when given an option to do so (Phinney and Alipuria 1996). Broad race/ethnic categories may be insufficient for launching appropriate health care interventions. For example, there is a difference of almost 50 percent in low birth weight rates and adequacy of prenatal care between Cambodians and American-born Chinese, a 25 percent difference between Cuban Americans and Puerto Ricans born in the continental United States, and about a 25 percent difference for mothers who identify themselves as Barbadian and Jamaican (National Research Council of the National Academies Workshop 2003). HCOs cannot understand the full diversity of the communities they serve and establish meaningful dialogues and partnerships without precise self-identification.

When an individual self-identifies as being from a certain population subgroup, it may also mean that the individual is more likely to have health

beliefs, health care use patterns, and perspectives about the health care system that are common to that community. Knowing that a person with diabetes is “white” may be less important than knowing that the person is from Poland in terms of understanding the attitudes, beliefs, and obstacles the patient may face while adapting to needed dietary and health behavior changes. Having this information enables HCOs to provide patient-centered care.

PRIMARY LANGUAGE

Limited-English speaking patients need to be able to communicate with their health care providers to ensure that the quality of their care is not compromised. Solid evidence has shown that language barriers can adversely affect quality of care (Timmins 2002; Ngo-Metzger 2003; Weech-Madonado et al. 2003). The 2000 U.S. Census supplementary survey indicates that 30.5 million U.S. residents were born in another country and foreign-born residents collectively speak over 150 languages. A significant number of people have limited English proficiency, that is, “a limited ability to read and write English and speak English “very well” (Ganey 2002). Obviously, questions about a patient’s language must be directly addressed to the individual because it is not possible for a registration or admitting clerk to “guess” a person’s language.

BARRIERS TO COLLECTING DATA DIRECTLY FROM PATIENTS

Although the rationales for collecting information directly from patients are numerous, there are some practical challenges that must be overcome to do this. Collecting information about the race and ethnicity of patients locally, where they go to get care—in hospitals, clinics, and physicians’ offices—can be sensitive for many reasons. Patients entering a hospital or doctor’s office expect that they are entering into a “caring relationship.” Asking questions about their race and ethnicity may undermine this relationship. Health plans can collect these data apart from where care is delivered (e.g., at time of enrollment), but there still may be concerns on the part of the patient/enrollee about (1) how the data will be used; (2) their privacy and methods to ensure confidentiality; and (3) how to answer questions if the categories provided are inadequate.

Negative Reactions from Patients

Within the therapeutic relationship, there is an element of unavoidable vulnerability and trust where the patient is at a disadvantage before caregivers (Heitman 1994). People who are knowledgeable and self-confident while healthy may feel powerless and uncertain as patients. Racial, ethnic, cultural, and linguistic discordance between patients and providers can increase the potential for misunderstanding, distrust, and ability to communicate effectively. It is therefore vital to collect information about patients' race, ethnicity, and language in a manner that does not erode trust or impair therapeutic relationships. Patients need to understand why they are being asked to provide this information and be given reassurances so that they feel comfortable doing so. We understand that even with reassurances, some patients may be put off by questions about their race/ethnicity. If an institution explains to patients that race/ethnicity data will be used to monitor quality, does this imply that health care disparities exist in the institution?

Concerns about alienating patients and fear of profiling are very real. We found that the vast majority of patients in a pilot study at the Northwestern Medical Faculty Foundation General Internal Medicine Clinic in Chicago agreed that it is important to collect data on race and ethnicity. Almost half stated that they were somewhat or very concerned that the data might be used to discriminate against patients. A small percentage even said they would be less likely to go to a hospital or clinic that collected race and ethnicity data. Thus, the study concluded that although patients see the importance of asking these questions, they need to understand the context of why they are asked to alleviate their concerns (Baker et al. 2005).

Lack of Categories That Fit the Patients' Perceptions

Most HCOs have collected race and ethnicity information by having patients choose from a limited list of categories. The current race and ethnicity categories used by the Office of Management and Budget (OMB) are the recommended standard and endorsed by the federal government. However, the OMB categories are also vague (Sondik, Madans, and Smith 2000). For example, people of Middle Eastern or Arab ancestry are often categorized or expected to self-identify as "white" or "Caucasian." This can be perplexing for individuals from Middle Eastern and Arab backgrounds who do not consider themselves white (Lopez 2002).

Self-identification is important for a number of reasons. For example, Caribbean and Latin immigrants often feel more akin to their national

identity than to a racial category defined by American society. The “Asian/Pacific Islander” and “Latino” categories homogenize immense heterogeneity among various nationalities (Flores and Moon 2002). For example people from the Philippines, Japan, Vietnam, Laos, and China are clumped under the broad category of “Asian/Pacific Islander” yet possess significant differences in relation to health beliefs, behaviors, and diets (Makimoto 1998).

So, collecting data using a defined set of choices may be the worst of two worlds: the information obtained is of limited value for representing patients self-perceptions, and the process of collecting the data may alienate patients who find that the categories offered do not match their self identity. The challenge posed, then, is how does one take these granular categories and use them for reporting or analytical purposes? Despite criticisms, the OMB categories can be used as the standard template into which the more granular self-reported descriptions can be “rolled up” or aggregated. However, HCOs should use the more granular information for providing patient-centered care and developing race/ethnicity specific quality initiatives (Baker et al. 2005, 2006). We discuss the challenges of rolling up or aggregating smaller categories into larger ones below.

Using Self-Reported Race/Ethnicity Information

The potential for problems associated with small numbers is a fair concern, however this should not pose a barrier for HCOs to target improvement in care delivery. We suggest that if a group with small sample sizes shows trends towards worse processes of care, it should be explored regardless of statistical significance. For example, in an on-going study by the authors, we have noted that Latino and black patients show similar trends in reduced time to percutaneous intervention (PCI) for acute myocardial infarction. Although we do not have sufficient numbers to detect statistical significance, we are reviewing their charts to better understand why this is happening. Is it due to disparities in care or is it due to some other reason? Further, we are currently examining racial/ethnic differences in quality of care for congestive heart failure and pneumonia. Although the number of Latino patients we are tracking is quite small, we conduct root-cause analyses of health care differences even though the values are not statistically significant.

For analytical purposes, the “small numbers” problem could be partially overcome by pooling data across multiple HCOs. Our pilot study results showed that “open-ended self-reported race/ethnicity” can be aggregated

when necessary with a high degree of accuracy. We found 93 percent agreement between the categories created using patients own words and the closed-ended responses (OMB categories). Finally, collecting race/ethnicity data using patients' own more specific terms resulted in lower rates of missing and unusable data compared with the standard OMB questions (Baker et al. 2005, 2006).

Time

Many hospitals and other HCOs indicate that they do not collect race, ethnicity, and language data because of costs. The time needed to collect and maintain the data were identified as major barriers in site visits to six leading hospitals in 2003 (Hasnain-Wynia, Pittman, and Pierce 2004). The most cited barrier during site visits were "time constraints during the registration process." However, our pilot study showed that the average time to completion, using an open-ended format which allowed patients or caretakers to self-report race/ethnicity, took an average of 37 seconds to complete (Baker et al. 2006).

Staff Discomfort

Interviews with hospital staff (e.g., admitting and registration clerks) indicate that staff too can feel uncomfortable asking patients questions about race and ethnicity, fearing that that they might be creating barriers to care by posing these questions. Some registration staff felt it inappropriate to ask these questions because they might lead to perceptions that racial minorities were treated differently. All in all, these specific points underscore a dominant concern—staff did not know how to interview patients in a consistent manner to get the right information while maintaining patient comfort with the interaction (Hasnain-Wynia, Pittman, and Pierce 2004).

OVERCOMING OBSTACLES

Increasing Patients' Comfort Level Providing Information about Their Race and Ethnicity

It is possible to increase how comfortable patients feel about providing information about their race, ethnicity, and language by explaining the reasons for collecting this information and how it will be used. In the pilot study at Northwestern Medical Faculty Foundation General Internal Medicine Clinic,

patients felt most comfortable providing the information when they were told that the information would be used to “monitor care to ensure that all patients get the best care possible” Three other statements were tested that said race and ethnicity data were being collected because: (1) government agencies require it, (2) it was needed to gain information to help hire and train staff, and (3) it would help make sure the patient got the best health care possible. These were less successful at improving patients’ comfort level, particularly for non-whites (Baker et al. 2005). Because some patients are concerned this information could be used to discriminate against them, HCOs need to assure patients that their confidentiality will be maintained and that access to the data will be highly restricted (American Medical Association 2003).

Addressing Staff Discomfort

If HCOs incorporate measures to address the concerns and discomfort of patients in providing information about their race and ethnicity, they will already have taken the most important step towards alleviating the concerns and discomfort of staff. Still, HCOs will need to invest in staff education and training to equip staff with the knowledge and tools they will need to implement a systematic method for collecting race, ethnicity, and primary language information from patients (Hasnain-Wynia and Pierce 2005). The main components of staff training should include:

- A script that outlines the rationale for collecting the information, how to ask the questions, and addresses how the information will (and will not) be used.
- Case examples to teach how staff should answer questions from patients who express concerns or who ask for more information about why this information is being collected.
- How to record/code the information.

Educational/training programs need to continue after initial implementation and should work to change organizational culture toward obtaining race, ethnicity, and language information from patients respectfully and efficiently. The success of such a program is dependent upon evaluation of its effectiveness and making appropriate changes on an on-going basis. HCOs should undertake community outreach efforts to explain why they are collecting this information and to initiate a dialogue about how it will be used.

PROPOSED ELEMENTS OF A UNIFORM FRAMEWORK AND CONCLUDING COMMENTS

In summary, we suggest that the elements of a uniform framework for collecting race, ethnicity, and language data in HCOs should include:

- (1) A rationale for why the patient is being asked to provide information about his/her race, ethnicity, and language.
- (2) A script for staff to use each time so that they ask questions in a uniform fashion.
- (3) A method for allowing patients to self-identify their race, ethnicity, and language using their own words rather than a preestablished set of categories.
- (4) A standardized approach for “rolling up” granular responses to the OMB categories for analytical and reporting purposes.
- (5) Assurances that the data will be held confidential and that a limited number of people will have access to the data, and a mechanism to guarantee this claim.

The uniform framework provides a process map for HCOs to systematically collect race, ethnicity, and language information from patients or their caregivers resulting in more accurate and complete data. HCOs with data from their own institutions may, in turn, use the information to reduce health care disparities, develop targeted initiatives to improve quality of care and provide patient-centered care.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of the Robert Wood Johnson Foundation in the development of this manuscript and the Commonwealth Fund for support of some of the key research discussed herein.

REFERENCES

- Agency for Healthcare Quality and Research. 2003. *The National Healthcare Disparities Report*. Washington, DC: AHRQ.
- America's Health Insurance Plans. 2004. Available at <http://www.ahip.org/content/fileviewr.aspx>

- American Medical Association, Institute for Ethics. 2003. *The Consensus Report on Privacy and Confidentiality*. Available at http://www.ama-assn.org/ama/upload/mm/369/ef_privacy_rpt.pdf
- Baker, DW., K. Cameron, J. Feinglass, P. Georgas, S. Foster, D. Pierce, J. Thompson, and R. Hasnain-Wynia. 2005. "Patients' Attitudes toward Health Care Providers Collecting Information about Their Race and Ethnicity." *Journal of General Internal Medicine*. 20 (10): 895–900.
- . 2006. "Developing and Testing of a System for Rapidly and Accurately Collecting Patients' Race and Ethnicity." *American Journal of Public Health*. 96 (3): 532–8.
- Bartlett, E. E., M. Grayson, R. Barker, D. M. Levine, A. Godlen, and S. Libber. 1984. "The Effects of Physician Communication Skills on Patient Satisfaction, Recall, and Adherence." *Journal of Chronic Diseases* 37: 755–64.
- Bocchino, C. 2004. "Racial and Ethnic Data Collection by Health Plans." In *Eliminating Disparities: Measurement and Data Needs*, National Research Council Report edited by M. Ver Plog and E. Perrin, pp. 272–87. Washington, DC: National Academies Press.
- Donohue, S. K. 2004. "General Counsel." Personal Communication, National Committee for Quality Assurance, August 4.
- Fiscella, K., P. Franks, M. Gold, and C. Clancy. 2000. "Inequality in Quality: Addressing Socioeconomic, Racial and Ethnic Disparities in Health Care." *Journal of the American Medical Association*. 283: 2579–84.
- Flores, L., and D. Moon. 2002. "Rethinking Race, Revealing Dilemmas: Imagining a New Racial Subject in Race Traitor." *Western Journal of Communication* 66 (2): 181–207.
- Fremont, A., and N. Lurie. 2004. "The Role of Racial and Ethnic Data Collection in Eliminating Disparities in Health Care." In *Eliminating Disparities: Measurement and Data Needs*, National Research Council Report edited by M. Ver Plog and E. Perrin, pp. 202–29. Washington, DC: National Academies Press.
- Ganey, F. 2002. Commonwealth Fund Grantee Meeting. Personal Communication, March 7.
- Hasnain-Wynia, R., and D. Pierce. 2005. "A Toolkit for Collecting Race, Ethnicity, and Primary Language Information from Patients." The Health Research and Educational Trust. Available at www.hret disparities.org
- Hasnain-Wynia, R., M. Pittman, and D. Pierce. 2004. "Who, When and How: The Current State of Race, Ethnicity, and Primary Language Data Collection in Hospitals." *The Commonwealth Fund*.
- Heitman, E. 1994. "Cultural Diversity and the Clinical Encounter: Intercultural Dialogue in Multi-Ethnic Patient Care." In *Theological Analysis of the Clinical Encounter*, edited by G. P. McKenny and J. R. Sande, pp. 203–23. Dordrecht: Kluwer Academic Publishers.
- Lopez, A. 2002. "Middle Eastern Populations in California: Estimates from the Census 2000 Supplementary Survey." *Center for Comparative Studies in Race and Ethnicity, Stanford University*.
- Lurie, N., M. Jung, and R. Lavizzo-Mourey. 2005. "Disparities and Quality Improvement: Federal Policy Levers." *Health Affairs* 4 (2): 354–64.

- Makimoto, K. 1998. "Drinking Patterns and Drinking Problems among Asian-Americans and Pacific Islanders." *Alcohol Health and Research World*. 22 (4): 270-5.
- Marshall, M., P. Shekelle, S. Leatherman, and R. Brook. 2000. "The Public Release of Performance Data." *Journal of the American Medical Association* 283: 1866-74.
- Melnick, D., and E. Perrin, (eds). 2003. "Improving Racial and Ethnic Data in Health: Report of a Workshop." *National Research Council*. Washington, DC: National Academies Press.
- Modernhealthcare.com. 2005. "Senators Propose Requiring Race Data from Hospitals" [accessed on October 28]. Available at <http://www.modernhealthcare.com/story.cms?newsid=4448&archive=N>
- Moscou, S., M. R. Anderson, J. B. Kaplan, and L. Valencia. 2003. "Validity of Racial/Ethnic Classifications in Medical Records Data: An Exploratory Study." *American Journal of Public Health* 93 (7): 1084-6.
- National Committee on Vital and Health Statistics Minutes. 2004. [Accessed on november 14, 2005]. Available at <http://www.ncvhs.hhs.gov/04061mn.htm>. June 16-17 .
- National Research Council of the National Academies. 2003. "Improving Racial and Ethnic Data in Health: Report of a Workshop." *Panel on DHHS Collection of Race and Ethnicity Data*. Washington, DC: National Academies Press.
- Nerenz, D. R., C. Currier, and K. Paez. 2004. "Collection of Data on Race/Ethnicity by Private Sector Organizations: Results of a Medical Group Survey." In *Eliminating Disparities: Measurement and Data Needs*, National Research Council Report edited by M. Ver Ploeg and E. Perrin, pp. 249-71. Washington, DC: National Academies Press.
- Nerenz, D. R., M. Gunter, M. Garcia, R. Green-Weir, K. Wisdom, and C. Joseph. 2002. "Developing a Health Plan Report Card on Quality of Care for Minority Populations." *The Commonwealth Fund*.
- Ngo-Metzger, Q., M. P. Massagli, B. R. Clarridge, M. Manocchia, R. Davis, L. Lezzoni, and R. Phillips. 2003. "Linguistic and Cultural Barriers to Care: Perspectives of Chinese and Vietnamese Immigrants." *Journal of General Internal Medicine* 18 (1): 44-52.
- Perot, R. T., and M. Youdelman. 2001. "Racial, Ethnic, and Primary Language Data Collection in the Health Care System: An Assessment of Federal Policies and Practices." *The Commonwealth Fund*.
- Phinney, J., and L. Alipuria. 1996. "At the Interface of Cultures: Multiethnic/Multi-racial High School and College Students." *Journal of Social Psychology* 136 (2): 139-58.
- Robert Wood Johnson Foundation. 2003. "Key Findings of a National Survey Conducted among Adults Who Have Health Care Coverage on Behalf of RWJF on the Issue of Disparities in Healthcare." Available at <http://www.rwjf.org/news/POSfullMemo.pdf>
- Sehgal, A. R. 2003. "Impact of Quality Improvement Efforts on Race and Sex Disparities in Hemodialysis." *Journal of the American Medical Association* 289 (8): 996-1000.

- Smith, T. W. 1997. "Measuring Race by Observation and Self-Identification." National Opinion Research Center, GSS Methodological Report, No. 89.
- Sondik, E. J., J. Madans, and S. Smith. 2000. "Race/Ethnicity and the 2000 Census: Implications for Public Health." *American Journal of Public Health* 90 (11): 1790–14.
- The American College of Physicians. 2004. "Position Paper: Racial and Ethnic Disparities in Health Care." *Annals of Internal Medicine* 141 (3): 226–32.
- Timmins, C. L. 2002. "The Impact of Language Barriers on Health Care of Latinos in the United States: A Review of the Literature and Guidelines for Practice." *Journal of Midwifery and Women's Health* 47 (2): 80–96.
- Torrassa, U. 2003. "Racial Initiative Ignites Medical Worries." *San Francisco Chronicle*, August 11.
- Ver Ploeg, M., and E. Perrins, 2004. (eds). *Eliminating Health Disparities: Measurement and Data Needs*. National Research Council. Washington, DC: National Academies Press.